

Symposium: Experimentelle Kompetenzen in den Naturwissenschaften (ExKoNawi)

International wie auch in der Schweiz gewinnt die an Kompetenzen orientierte Bildungssteuerung und -evaluation zunehmend an Bedeutung. Um zu überprüfen, inwieweit der naturwissenschaftliche Unterricht die in Lehrplänen und Bildungsstandards beschriebenen experimentellen Kompetenzen bei Schülerinnen und Schülern fördert, werden sowohl für die Individualdiagnostik als auch für large-scale Assessments entsprechende Experimentiertests benötigt. Insbesondere sind hands-on Assessments gefragt, bei denen die Schülerinnen und Schüler beim Lösen der Aufgaben einen authentischen Experimentier- und Erkenntnisgewinnungsprozess durchlaufen (z. B. Shavelson, Baxter & Pine, 1991). Gegenüber Papier- und Bleistift-Tests weisen solche Tests zum einen eine erhöhte Inhaltsvalidität auf, zum anderen können sie als Vorbild für eine zeitgemäße, konstruktivistisch ausgerichtete Experimentierpraxis im Unterricht dienen (z. B. Aschbacher, 1991; Stebler, Reusser & Ramseier, 1998). Zur Erfassung fächerübergreifender experimenteller Kompetenzen, wie sie in der Schweiz im integrierten Naturwissenschaftsunterricht der Volksschule im Vordergrund stehen, fehlen bisher jedoch ausreichend reliable und valide Performanztests. Dies liegt unter anderem daran, dass theoriebasierte Konzepte fehlen, um Kompetenzprogressionen a priori zu modellieren und so die Konstruktion anforderungshomogener Experimentieraufgaben sicherzustellen. Spezifisch für interdisziplinäre Experimentiertests wird zudem eine Hauptursache in den zugrundeliegenden strukturellen Kompetenzmodellierungen vermutet, die sich bisher zu wenig an den für die Aufgabenlösung erforderlichen, über Aufgaben zu verschiedenen naturwissenschaftlichen Fachkontexten transferierbaren Problemlösestrategien orientieren. Außerdem schränken messtechnische Probleme die Validität von Experimentiertests zusätzlich ein (vgl. Gut, 2012; Gut et al., 2014; Theyssen et al., 2014).

Kompetenzmodellierung bei ExKoNawi

Im Rahmen des Projektes *ExKoNawi* (Experimentelle Kompetenzen in den Naturwissenschaften) wurde ein interdisziplinäres Kompetenzstrukturmodell entwickelt, das auf dem Problemtypenansatz (z. B. Gott & Duggan, 1995; Shavelson & Ruiz-Primo, 1999) beruht. Der Problemtypenansatz ermöglicht, die verschiedenen experimentellen Aktivitäten in der Schulpraxis fächerübergreifend und curriculumsvalide zu beschreiben und deren Progressionen a priori zu modellieren, indem diese Aktivitäten nach ihren Problemstellungen differenziert und idealisierten Problemtypen zugeordnet werden. In diesem Sinne wurden auf der Basis des Lehrplans und aufgrund einer Analyse bestehender hands-on-Tests (z. B. Archenhold, 1988; Schofield, 1989; Erickson et al., 1992; Harmon et al., 1997; Shavelson & Ruiz-Primo, 1999; Gut, 2012) verschiedene idealisierte Problemstellungen unterschieden, von denen zunächst vier Problemtypen modelliert wurden (siehe Tab. 1): kategoriengeleitetes Beobachten (vgl. Stevens, 1978; Gott & Welford, 1987; Solano-Flores, Shavelson, Ruiz-Primo et al., 1997; Hammann, 2002), skalenbasiertes Messen (Millar et al., 1994; Lubben & Millar, 1996; Masnick & Klahr, 2003; Lippman Kung, 2005; Volkwyn, Allie & Buffler, 2008; Buffler, Lubben & Ibrahim, 2009; Heinicke, Glomski, Priemer & Riess, 2010; Munier, Merle & Brehelin, 2011; Hellwig, 2012), fragengeleitetes Untersuchen (Kuhn & Phelps, 1982; Linn, Clement & Pulos, 1983; Donnelly, 1987; Schauble, Glaser, Raghavan & Reiner, 1991; Schauble, Klopfer & Raghavan, 1991; Schauble, Glaser, Raghavan & Reiner, 1992; Song & Black, 1992; Kanari & Millar, 2004; Hammann, Phan, Ehmer & Bayrhuber,

2006; Furtak, Seidel, Iverson & Briggs, 2012; Schwichow, Croker, Zimmerman et al., 2016) und effektbasiertes Vergleichen (Erickson, 1994; Solano-Flores, 1994; Meyer & Carlisle, 1996).

Problemtyp	Beschreibung
kategoriengeleitetes Beobachten	Phänomene anhand gegebener Kategorien (Begriffe) beschreiben und vergleichen
skalenbasiertes Messen	quantitative Größen mit gegebenen Messinstrumenten (Skala) genau messen
fragengeleitetes Untersuchen	korrelative Zusammenhänge zwischen gegebenen Variablen (Frage) untersuchen
effektbasiertes Vergleichen	Objekte anhand einer gegebenen Eigenschaft experimentell (ohne direkte Messung der Eigenschaft) vergleichen

Tab. 1: Strukturmodellierung von ExKoNawi

Das Modell kann prinzipiell durch Hinzunahme weiterer bereits angedachter Problemtypen, wie beispielsweise funktionsgeleitetes Konstruieren, effektbasiertes Testen oder kriteriengeleitetes Klassifizieren, erweitert werden. Die Kompetenzprogression wird für jeden Problemtyp separat modelliert, wobei maximal fünf Leistungsniveaus unterschieden werden. Mit typenspezifischen Progressionsmodellen können spezifische, aus der Theorie und Erfahrung hergeleitete Qualitätsstandards einheitlich bewertet werden. Post hoc kann zudem überprüft werden, ob eine empirische Schwierigkeitsprogression der verschiedenen Qualitätsstandards festgelegt werden kann (Progressionsmodellierung). Experimentieraufgaben desselben Problemtyps sind so über die Kontexte hinweg vergleichbar.

ExKoNawi-Experimentiertest

Passend zum Kompetenzmodell wurden für jeden der oben beschriebenen Problemtypen je zwei Experimentieraufgaben zu einem biologischen, chemischen und physikalischen Kontext entwickelt. Alle Aufgaben sind ausschließlich hands-on Experimente für die Sekundarstufe I, welche von den Schülerinnen und Schülern in 18 Minuten Einzelarbeit zu lösen sind. Um die kompetenzirrelevanten Aufgabenanforderungen möglichst konstant zu halten, wurde auf größtmögliche Homogenität der schriftlichen Aufgabenstellungen und Antwortformate innerhalb der Problemtypen geachtet. Um den Einfluss von kontextspezifischem Fachwissen zu minimieren, wurden vorwiegend Alltagskontexte und alltagssprachliche Formulierungen verwendet. Die Aufgaben bestehen je nach Problemtyp aus zwei bis drei Teilaufgaben (Fragen), die meist aufbauend gelöst werden (siehe **Fehler! Verweisquelle konnte nicht gefunden werden.**).

	kategoriengeleitetes Beobachten	skalenbasiertes Messen	fragengeleitetes Untersuchen	effektbasiertes Vergleichen
Teilaufgaben	<ul style="list-style-type: none"> ein einzelnes Phänomen beschreiben Unterschiede und Gemeinsamkeiten zweier Phänomene beschreiben 	<ul style="list-style-type: none"> adäquates Messinstrument wählen (um die Genauigkeit zu erhöhen) eine quantitative Größe möglichst genau messen Messwiederholungen berücksichtigen (um die Genauigkeit zu erhöhen) 	<ul style="list-style-type: none"> einen gegebenen Zusammenhang untersuchen einen zweiten Zusammenhang unter Berücksichtigung des ersten untersuchen 	<ul style="list-style-type: none"> zwei Objekte anhand einer gegebenen Eigenschaft qualitativ vergleichen drittes Objekt qualitativ mit vergleichen alle drei Objekte quantitativ vergleichen, um Differenzen abzuschätzen

Tabelle 2: Aufgabenkonstruktion: Abfolge der Teilaufgaben/Inhalt der Fragestellungen

Die Ergebnissicherung erfolgt mittels Eigenrapportierung in gedruckte Testhefte. Dabei werden die Jugendlichen zunächst jeweils aufgefordert, ihre Beobachtungen, Messungen, Untersuchungen oder Vergleiche zu protokollieren. Diese Protokolle werden durch explizite Fragen nach den Resultaten, Reflexionen über die Durchführung und das erhaltene Ergebnis sowie abschließende Kontrollfragen ergänzt. Jede Experimentieraufgabe wird als ein Item kodiert (Vermeidung lokaler Itemabhängigkeiten), wobei je nach Problemtyp drei bis fünf Qualitätsstandards bewertet werden. Jeder Qualitätsstandard setzt sich in der Regel aus mehreren dichotom kodierten Kriterien zusammen, zu deren Beurteilung die Antworten zu allen Teilaufgaben berücksichtigt werden. Ein Qualitätsstandard ist erreicht, wenn eine a priori normativ festgelegte Zahl der Kriterien (zwischen der Hälfte und zwei Drittel der möglichen Punktzahl) erfüllt sind.

Pilotierung der Experimentieraufgaben und Validitätsprüfungen

Nachdem zwei Teilvalidierungen mit je drei Experimentieraufgaben zu je zwei Problemtypen durchgeführt wurden (Gut et al., 2014; Hild et al., 2015; Metzger et al., 2014), wurde das Modell mit weiteren je drei Experimentieraufgaben pro Problemtyp validiert (Gut et al., in diesem Band). Zusätzlich wurde die Konstruktvalidität für „effektbasiertes Vergleichen“ geprüft (Hild et al., in diesem Band). Anschließend wurde der ExKoNawi-Experimentiertest so überarbeitet, dass er nun in einer groß angelegten Studie validiert werden kann (Bonetti et al., in diesem Band).

Literatur

- Archenhold, F. (Hrsg.). (1988). *Science at age 15: a review of APU survey findings 1980-84*. London: Her Majesty's Stationery Office.
- Aschbacher, P. R. (1991). Performance assessment: state activity, interest, and concerns. *Applied Measurement in Education*, 4(4), 275-288.
- Buffler, A., Lubben, F. & Ibrahim, B. (2009). The relationship between students' views of the nature of science and their views of the nature of scientific measurement. *International Education Journal*, 31(9), 1137-1156.
- Donnelly, J. F. (1987). Fifteen-year-old pupils' variable handling performance in the context of scientific investigations. *Research in Science & Technological Education*, 5(2), 135-147.
- Erickson, G. (1994). Pupil's understanding of magnetism in a practical assessment context: the relationship between content, process and progression. In P. Fensham, G. Richard & R. White (Hrsg.), *The content of science* (S. 80-97). London: Falmer.
- Erickson, G., Bartley, A. W., Blake, L., Carlisle, R. W., Meyer, K. & Stavy, R. (1992). *British Columbia assessment of science 1991 technical report II: Student performance component*. Victoria: Queen's Printer.
- Furtak, E. M., Seidel, T., Iverson, H. & Briggs, D. C. (2012). Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching: A Meta-Analysis *Review of Educational Research*, 82(3), 300-329.
- Gott, R. & Duggan, S. (1995). *Investigative work in the science curriculum*. Buckingham: Open University Press.
- Gott, R. & Welford, G. (1987). The assessment of observation in science. *School Science Review*, 69, 217-227.
- Gut, C. (2012). *Modellierung und Messung experimenteller Kompetenz - Analyse eines large-scale Experimentiertests*. Berlin: Logos.
- Gut, C., Metzger, S., Hild, P. & Tardent, J. (2014). Problemtypenbasierte Modellierung und Messung experimenteller Kompetenzen von 12- bis 15-jährigen Jugendlichen. *PhyDid B, Didaktik der Physik, Beiträge zur DPG-Frühjahrstagung*, 9 S.
- Hammann, M. (2002). *Kriteriengeleitetes Vergleichen im Biologieunterricht*. Innsbruck: StudienVerlag.
- Hammann, M., Phan, T. T. H., Ehmer, M. & Grimm, T. (2008). Assessing pupil's skills in experimentation. *Journal of Biological Education*, 42(2), 66-72.
- Harmon, M., Smith, T. A., Martin, M. O., Kelly, D. L., Beaton, A. E., Mullis, I. V. S., . . . Orpwood, G. (1997). *Performance Assessment in IEA's Third International Mathematics and Science Study*. Chestnut Hill: Boston College.
- Heinicke, S., Glomski, J., Priemer, B. & Riess, F. (2010). Aus Fehlern wird man klug. Über die Relevanz eines adäquaten Verständnisses von "Messfehlern". *Praxis der Naturwissenschaften - Physik in der Schule*, 59(5), 26-33.

- Hellwig, J. (2012). *Messunsicherheiten verstehen - Entwicklung eines normativen Sachstrukturmodells am Beispiel des Unterrichtsfaches Physik*. Ruhr-Universität Bochum, Bochum. Zugriff unter <http://www-brs.ruhr-uni-bochum.de/netahtml/HSS/Diss/HellwigJulia/diss.pdf>
- Hild, P., Tardent, J., Gut, C. & Metzger, S. (2015). Projekt ExKoNawi: Typenspezifische Kompetenzprogressionen bei hands-on Testaufgaben. In: Bernholt, Sascha (Hrsg.). *Heterogenität und Diversität – Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht*, 145-147.
- Kanari, Z. & Millar, R. (2004). Reasoning from data: how students collect and interpret data in science investigations. *Journal of Research in Science Teaching*, 41(7), 748-769.
- Kuhn, D. & Phelps, E. (1982). The development of problem-solving strategies. *Advances in Child Development and Behavior*, 17, 1-44.
- Linn, M. C., Clement, C. & Pulos, S. (1983). Is it formal if it's not physics? The influence of content on formal reasoning. *Journal of Research in Science Teaching*, 20(8), 755-770.
- Lippman Kung, R. (2005). Teaching the concepts of measurement: an example of a concept-based laboratory course. *American Journal of Physics*, 73(8), 771-777.
- Lubben, F. & Millar, R. (1996). Children's ideas about the reliability of experimental data. *International Journal of Science Education*, 18(8), 955-968.
- Masnack, A. M. & Klahr, D. (2003). Error matters: an initial exploration of elementary school children's understanding of experimental error. *Journal of Cognition and Development*, 4(1), 67-98.
- Metzger, S., Gut, C., Hild, P. & Tardent, J. (2014). *Modelling and assessing experimental competence: an interdisciplinary progress model for hands-on assessments*. Konferenzbeitrag ESERA, Nikosia.
- Meyer, K. & Carlisle, R. W. (1996). Children as experimenters. *International Journal of Science Education*, 18(2), 231-248.
- Millar, R., Lubben, F., Gott, R. & Duggan, S. (1994). Investigating in the school science laboratory: conceptual and procedural knowledge and their influence on performance. *Research Papers in Education*, 9(2), 207-248.
- Munier, V., Merle, H. & Brehelin, D. (2011). Teaching scientific measurement and uncertainty in elementary school. *International Journal of Science Education*, 1-32, iFirst Article.
- Schauble, L., Glaser, R., Raghavan, K. & Reiner, M. (1991). Causal models and experimentation strategies in scientific reasoning. *The Journal of the Learning Sciences*, 1(2), 201-238.
- Schauble, L., Glaser, R., Raghavan, K. & Reiner, M. (1992). The integration of knowledge and experimentation strategies in understanding a physical system. *Applied Cognitive Psychology*, 6, 321-343.
- Schauble, L., Klopfer, L. E. & Raghavan, K. (1991). Students' transition from an engineering model to a science model of experimentation. *Journal of Research in Science Teaching*, 28(9), 859-882.
- Schofield, B. (Hrsg.). (1989). *Science at age 13: a review of APU survey findings 1980-84*. London: Her Majesty's Stationery Office.
- Schwichow, M., Croker, S., Zimmerman, C., Höffler, T. & Härtig, H. (2016). Teaching the control-of-variables strategy: a meta-analysis. *Developmental Review* 39, 37-63.
- Shavelson, R. J., Baxter, G. P. & Pine, J. (1991). Performance assessment in science. *Applied Measurement in Education*, 4(4), 347-362.
- Shavelson, R. J. & Ruiz-Primo, M. A. (1999). Leistungsbewertung im naturwissenschaftlichen Unterricht. *Unterrichtswissenschaft*, 27(2), 102-127.
- Solano-Flores, G. (1994). *A logical model for the development of science performance assessments*. University of California, Santa Barbara. Solano-Flores, G., Shavelson, R. J., Ruiz-Primo, M. A., Schultz, S. E. & Wiley, E. W. (1997). On the development and scoring of classification and observation science performance assessments. *CSE Technical Report* (Vol. 458). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.
- Song, J. & Black, P., J. (1992). The effects of concept requirements and task contexts on pupils' performance in control of variables. *International Journal of Science Education*, 14(1), 83-93.
- Stebler, R., Reusser, K. & Ramseier, E. (1998). Praktische Anwendungsaufgaben zur integrierten Förderung formaler und materialer Kompetenzen. *Bildungsforschung und Bildungspraxis*, 20(1), 28-53.
- Stevens, P. (1978). On the Nuffield philosophy of science. *Journal of Philosophy of Education*, 12, 99-111.
- Theysen, H., Schecker, H., Gut, C., Hopf, M., Kuhn, J., Labudde, P., . . . Vogt, P. (2014). Modelling and assessing experimental competencies in physics. In C. Bruguère, A. Thibergien & P. Clément (Hrsg.), *Topics and trends in current science education. 9th ESERA conference selected contributions* (S. 321-338). Dordrecht: Springer.
- Volkwyn, T. S., Allie, S. & Buffler, A. (2008). Impact of conventional introductory laboratory course on the understanding of measurement. *Physical Review Special Topics - Physical Education Research*, 4, 1-10.